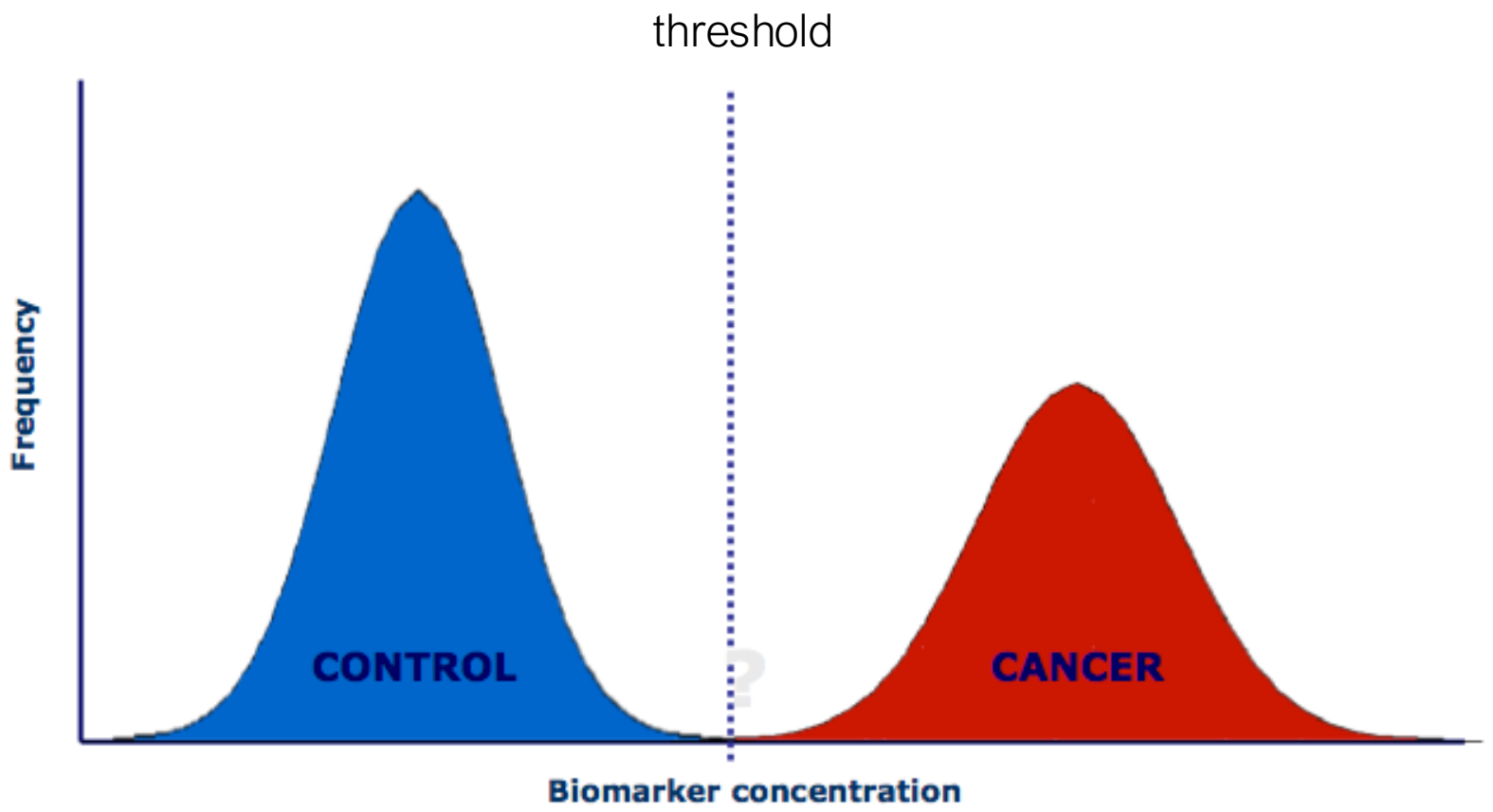
(6) Binary Diagnostic tests (screening assays/dichotomous diagnostic tests)

In the field of diagnostic tests and analytical tests with binary classification of the output (e.g.: pregnancy test, biomarker tests, antibody tests), the terms "Sensitivity" and "Specificity" have a peculiar meaning.
The binary tests' outputs is either positive (ie: concentration of the biomarker or other target molecule) is above a certain threshold) or negative (ie: concentration of the biomarker or other target molecule) is below a certain threshold.
"Sensitivity": probability that the test provides a correct answer in the occurrence that the sample is "positive".
"Specificity": probability that the test provides a correct answer in the occurrence that the sample is "negative".

These two parameters indicate the performance of the assay and they depend on:

- the relevance of the biomarker or analyte molecule measured (typically a molecule in a fluid or tissue, sometimes a physical parameter) employed to classify the population in two groups,
- the choice of the "threshold concentration" (synonyms: biomarker concentration, criterion value, cut-point, cut-off, criterion value) of the biomarker.

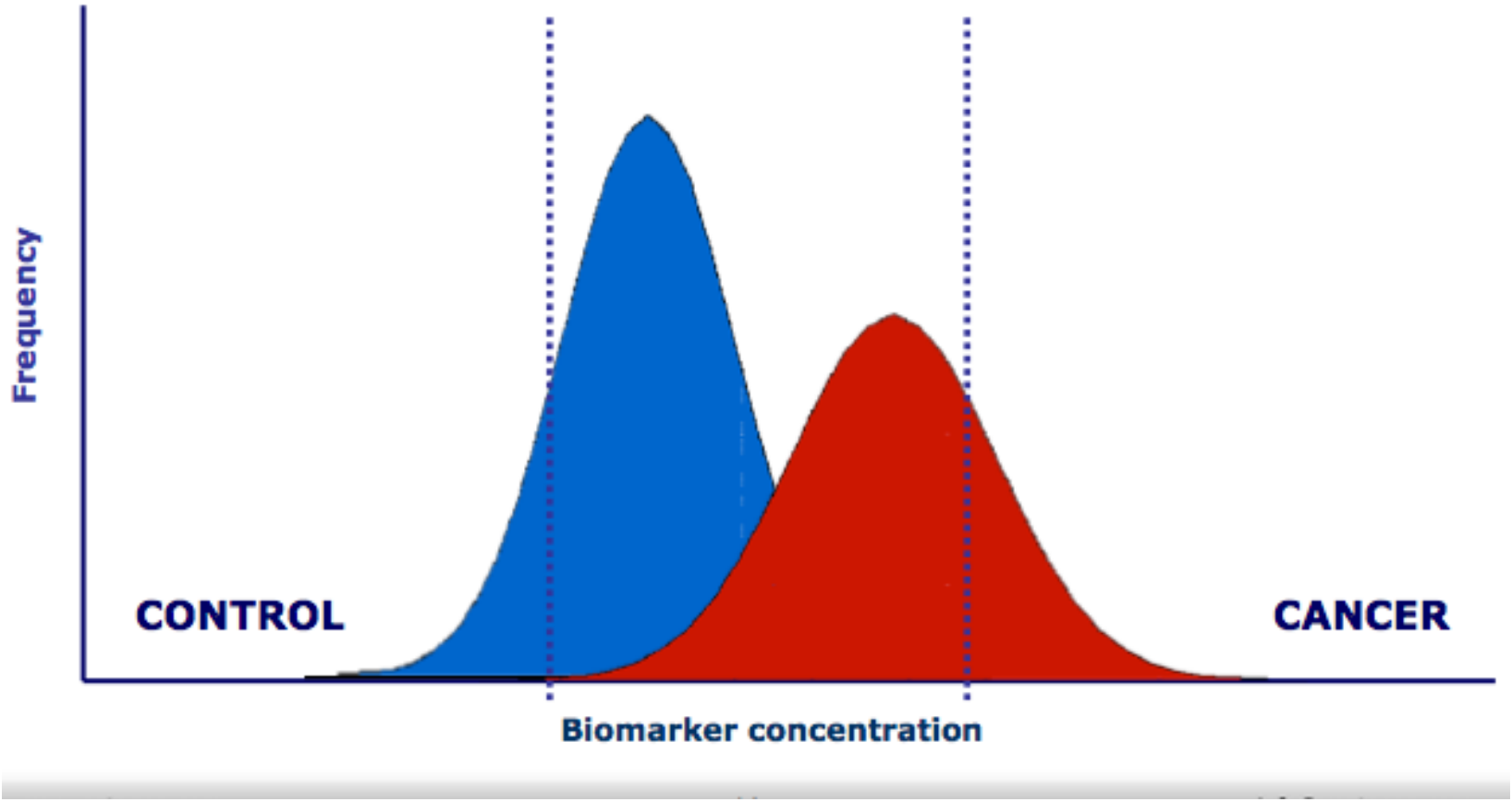
The biomarker concentration is perfectly correlated with the patient condition.



Certain biomarker concentration values can be observed both for healthy and sick subjects.

threshold that grants high sensitivity

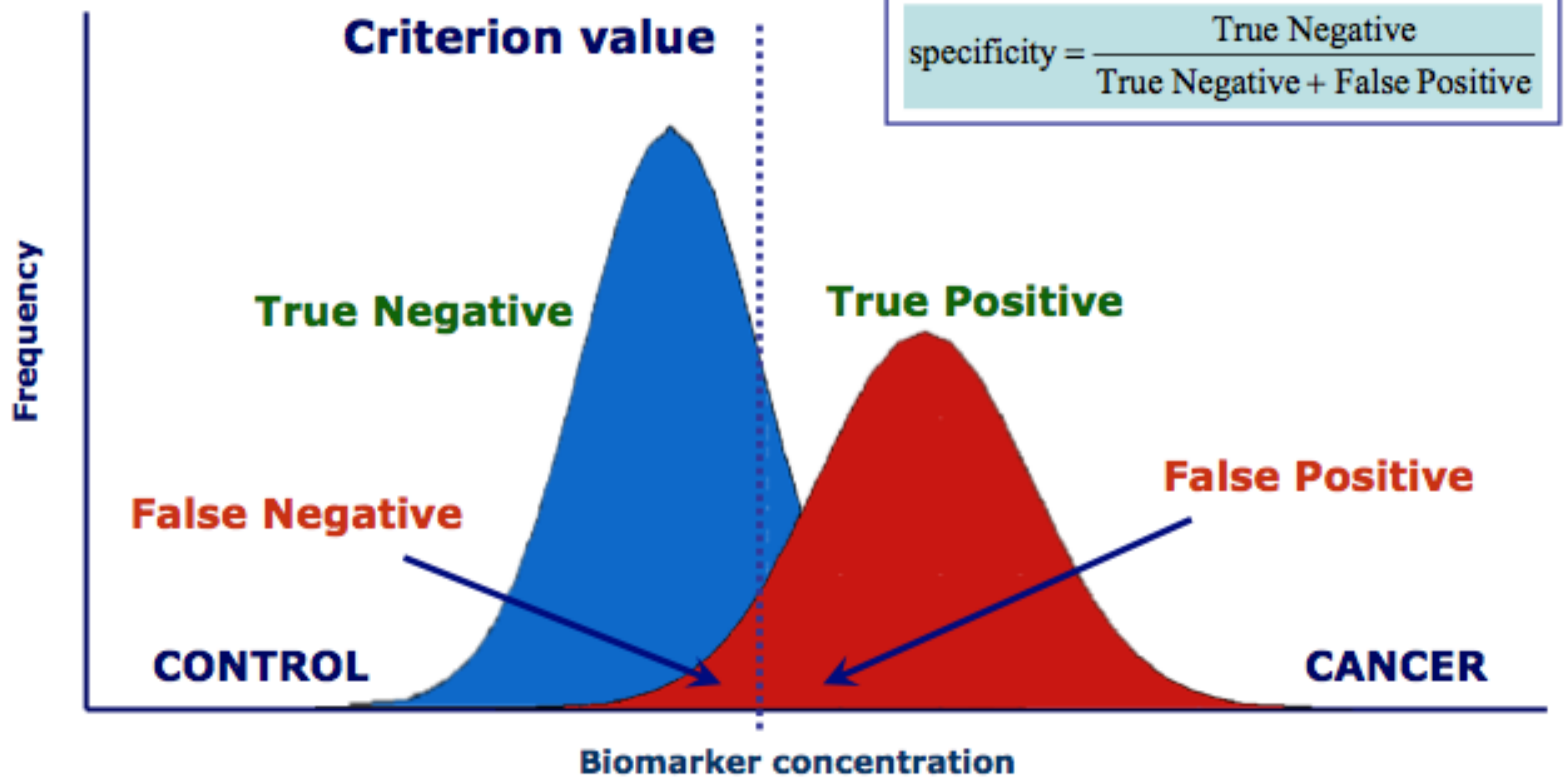
threshold that grants high specificity



Sensitivity and Specificity

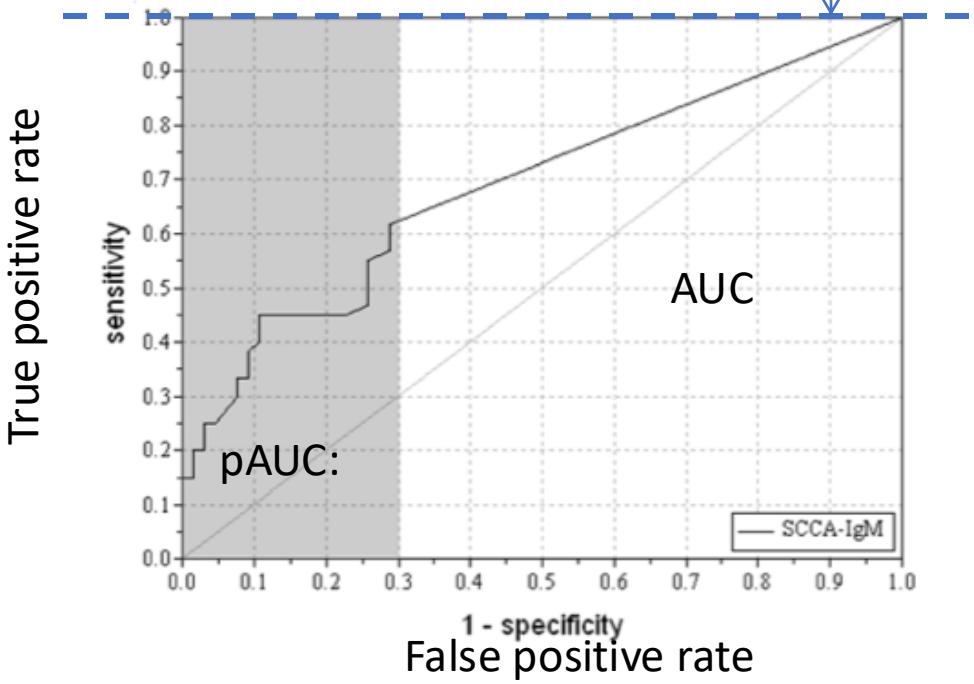
$$\text{sensitivity} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}$$

$$\text{specificity} = \frac{\text{True Negative}}{\text{True Negative} + \text{False Positive}}$$



Receiver operating characteristic curve

The ROC curve reaches this level if negative patients has lower biomarker values than any positive patient.

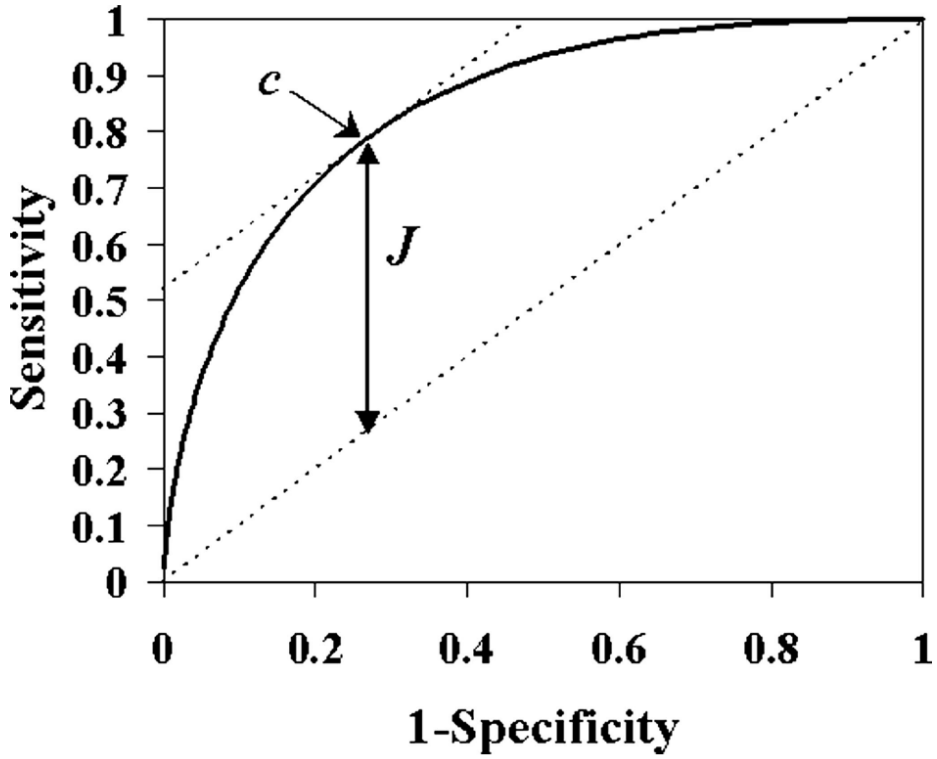


AUC: Area under the ROC curve
-to compare diagnostic tests
-pAUC: partial AUC: restricted region.

<u>Different biomarkers can be compared based on the value of the pAUC for a given specificity</u>.

ROC curve from 60 patients with HCC (hepatocellular carcinoma) versus 66 donors. AUC=0.689, pAUC=0.121.

Youden index J (first formalized by C. S. Peirce 1884 in Science, then applied to diagnostic tests by W.J. Youden in 1950



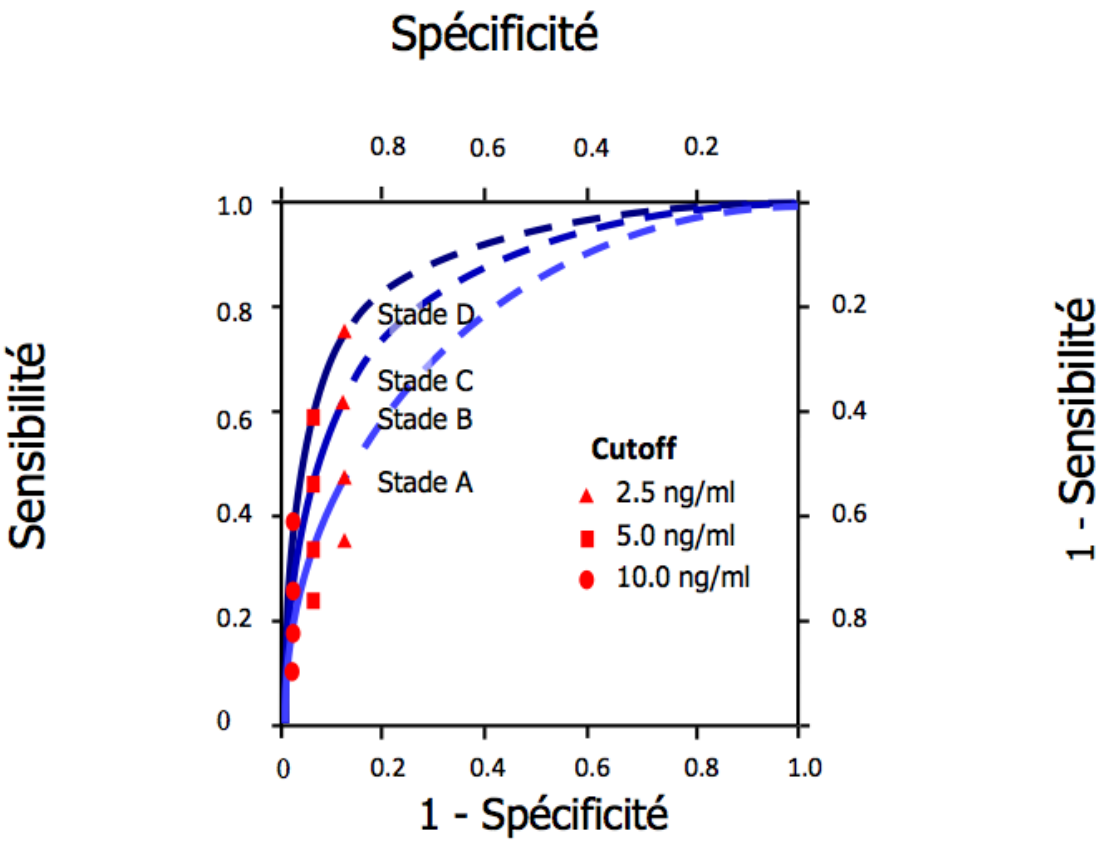
c optimal cut-off
diagonal line is the chance line:
J defines the best choice of the cut-off for this ROC curve

J = maximum {sensitivity (c) + specificity p(c) - 1}

J occurs at the optimal cut-point for calling a patient diseased, maximizing the number of correctly classified individuals.
On the other hand, the consequences of a positive or negative test result (ie, intervention) may be quite different and the loss from missing a case may be greater than from overcalling a control.

Diagnostic test

Test exhibits different performance depending on the diesease stage



Policlinique Médicale Universitaire CH-Lausanne

Precision

PRECISION: The probability that a patient has cancer if the test result is positive: p(cancer|positive)

This depends not only on the performance of the test but also on the **prevalence** of the disease in the population. The prevalence is p(cancer)

Bayes' theorem:
p(cancer|positive)= p(cancer) * p(positive|cancer)/p(positive)
where

p(positive|cancer) is the conditional probability that the test would be positive in case of patient with cancer: Sensitivity
p(positive): probability that the test is positive (sum of false positives and true positives)

Precision can be also calculated as:

PRECISION = True positives /(false positives + true positives)

And as

PRECISION = prevalence*sensitivity / [(prevalence * sensitivity) + [(1-Prevalence) * (1 - Specificity)]]

The PRECISION reveals the importance of high SPECIFICITY

If prevalence is 1%, let's calculate the precision for the following cases:

- Sensitivity =0.7 and Specificity = 1; Precision=1
- Sensitivity =0.7 and Specificity = 0.95; Precision=0.12

- Sensitivity =1 and Specificity = 0.95; Precision=0.17
- Sensitivity =1 and Specificity = 0.7; Precision=0.03

- Sensitivity =0.6 and Specificity = 0.7; Precision=0.02
- Sensitivity =1 and Specificity = 0.7; Precision=0.03

If prevalence is 30%, let's calculate the precision for the following cases:

- Sensitivity =1 and Specificity = 0.95; Precision=0.9
- Sensitivity =1 and Specificity = 0.7; Precision=0.59
- Sensitivity =0.6 and Specificity = 0.7; Precision=0.46

- Sensitivity =0.7 and Specificity = 1; Precision=1
- Sensitivity =0.7 and Specificity = 0.95; Precision=0.86

PPV and NPV

Positive predictive value:

It is the ratio of patients truly diagnosed as positive to all those who had positive test results (including healthy subjects who were incorrectly diagnosed as patient). This characteristic can predict how likely it is for someone to truly be patient, in case of a positive test result.

Positive predictive value = $TP/(TP+FP)$

Negative predictive value:

It is the ratio of subjects truly diagnosed as negative to all those who had negative test results (including patients who were incorrectly diagnosed as healthy). This characteristic can predict how likely it is for someone to truly be healthy, in case of a negative test result.

Negative predictive value=$TN/(TN+FN)$

Predictive values and the prevalence of the disease

Since the ratio PPV and NPV includes both healthy and patient subjects, predictive values are affected by the prevalence of the disease and can differ from one setting to another for the same diagnostic test.

The lower the prevalence of the disease, the higher its negative predictive value.

On the other hand, the higher the prevalence of the disease, the higher the positive predictive value.

Example 1: Imagine we have a sample population of 100 people, 50 healthy and the others are patients. If the test was positive for 75 people of this population, the PPV and NPV of test are as follows (assuming specificity 1):

PPV: 50/75 = 0.66 or 66.6%. This means that in this population, 66.6% of people whose test result is positive, have the disease.

NPV: 25/25 = 100%. This means that in this population, 100% of the people whose test result is negative, are healthy